Feedback visualization in a grammar-based e-learning system for German: a preliminary user evaluation with the COMPASS system

Karin Harbusch[1] and Annette Hausdörfer[2]

Abstract. COMPASS[3] is an e-learning system that can visualize grammar errors during sentence production in German as a first or second language. Via drag-and-drop dialogues, it allows users to freely select word forms from a lexicon and to combine them into phrases and sentences. The system's core component is a natural-language generator that, for every new word the user wishes to attach to the current string (as an extension of this string or as a replacement of a substring), checks whether this tentative attachment is grammatically well-formed or not. On this basis, the system can compute and display online the grammatical structure of input strings in the form of syntactic trees, and identify and diagnose input errors. In the following, we focus on the crucial question of how to present the feedback to the learner. We propose tutored visualizations with animations of pedagogical agents. We briefly report the results of a preliminary user evaluation study in which the participants judged the well-formedness of prefabricated input sentences. The data, collected by means of eye-tracking and a questionnaire, show that L1 learners who are exercising an unfamiliar and error-prone grammatical structure, pay due attention to, and can profit from, this type of visualized error feedback.

Keywords: ICALL, grammar teaching, natural-language generation, personalized feedback, evaluation.

1. Introduction

Automatically generating personalized, reliable and immediate learner feedback is an important prerequisite for effective learning in Intelligent Computer-Assisted

1. University of Koblenz-Landau, Koblenz, Germany; harbusch@uni-koblenz.de
2. University of Koblenz-Landau, Koblenz, Germany; ahausdrfer@uni-koblenz.de
3. COMPASS stands for COMbinatorial and Paraphrastic Assembly of Sentence Structure (for more information see: https://userpages.uni-koblenz.de/~harbusch/COMPASS/index.html).

How to cite this article: Harbusch, K., & Hausdörfer, A. (2016). Feedback visualization in a grammar-based e-learning system for German: a preliminary user evaluation with the COMPASS system. In S. Papadima-Sophocleous, L. Bradley & S. Thouësny (Eds), *CALL communities and culture – Short papers from EUROCALL 2016* (pp. 178-184). Dublin: Research-publishing.net. https://doi.org/10.14705/rpnet.2016.eurocall2016.558

Language-Learning (ICALL). Molloy and Boud (2014) point out that "[f]eedback is a key process in learning, providing information on actual performance in relation to the goal of performance. […] There is mounting survey data to suggest that students are dissatisfied with feedback" (pp. 413-414); see Narciss (2008) for a thorough overview of feedback strategies in language learning. In the area of German as a second language, studies by Diehl et al. (2000; native speakers of French at various high school levels) and Ballestracci (2005; native speakers of Italian at university level) underscore the importance of feedback tuned to the learner's current performance level and his/her understanding of the explicitly presented grammar rules (see also Kartchava, 2012; van der Kleij, 2013; Varnosfadrani & Ansari, 2011).

In this paper, we focus on automatic feedback presentation in COMPASS, an e-learning system for German as a first and second language (described in several papers by Harbusch & Kempen, 2011; Harbusch, Härtel, & Cameran, 2013; Harbusch, Cameran, & Härtel, 2014). It confronts the learners only with errors they could have avoided, given their current proficiency level and their current understanding of the grammar rules. The relevant feedback calculated according to the learner's proficiency level is provided by an animated tutor pointing out which problem COMPASS has identified and where it is located in the syntactic structure. In a user study we explore how the feedback visualization is perceived by test subjects.

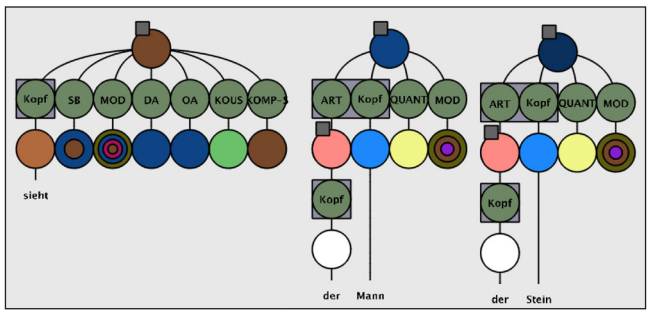
In Section 2, we argue for tutored visualizations with an animated pedagogical agent. Section 3 describes a preliminary user study with L1 learners as participants who judged the well-formedness of prefab input sentences that provided unfamiliar and error-prone grammatical structures. In the final Section 4, we draw some conclusions and address future work.

2. Feedback in COMPASS

In COMPASS, the user selects word forms from a lexicon and combines them into phrases and sentences via drag-and-drop dialogues. In response to each user action, the system displays the grammatical structure of the resulting word-form string in the form of a syntactic tree (see Figure 1). The system provides feedback on the (un)grammaticality of the string. In the example, both noun phrases are syntactically licensed as subject, although, 'der Stein' is semantically not an appropriate actor. However, none of the two can become the direct (cf. grammatical

function OA[4]=accusative object) or indirect object (cf. DA(=dative object) node) of the verb 'sieht' due to case mismatch.

Figure 1. A snapshot of the COMPASS workspace illustrating a scenario where the learner has selected the verb form 'sieht' *sees*. Moreover (s)he has assembled and word ordered the two noun phrases 'der Mann' *the man* and 'der Stein' *the stone*. (N.B. linearization checking is activated by drawing nodes into the grey boxes around grammatical functions provided by each Kopf/head; outside the boxes only the hierarchical structure – relations between nodes – is validated)



As mentioned by Harbusch et al. (2013),

> "the grammar formalism underlying COMPASS is *Performance Grammar* (Kempen & Harbusch, 2002), which uses separate rules for the hierarchical *structure* of a sentence and the *linear order* of its constituents. This split allows the student to divide a sentence construction exercise into relatively small parts. For instance, the learner can select a word, and inflect it as required by the intended grammatical function, without simultaneously considering the linear position of the constituent in the sentence under construction. At any time during this 'scaffolded' sentence construction process, the syntactic tree built so far remains visible on the screen, ready to be expanded or modified with additional words or phrases. Any sentence construction step can be undone and replaced online" (p. 105).

Another advantage arises from the generation-based approach of COMPASS, where the learner and system assemble an unambiguous syntactic tree together. Compared to other systems that allow free sentence construction by the user but

4. The annotations in trees resemble the ones in the TIGER corpus (Brants et al., 2004; e.g. KOUS=Unterordnende Konjunktion /subordinating conjunction).

rely on natural-language parsing, the feedback in a generation-based system is not hampered by ambiguous syntactic structures, meaning that feedback can be precise and exhaustive. However, as outlined in Section 1, it is desirable to highlight only those errors against rules the student is supposed to have mastered. COMPASS uses underspecification of grammar rules to accomplish this (cf. Harbusch et al., 2014; in the example in Figure 1, beginners might attach any of the two NPs as SB/OA/DA as case agreement is 'overlooked' by COMPASS on that proficiency level).

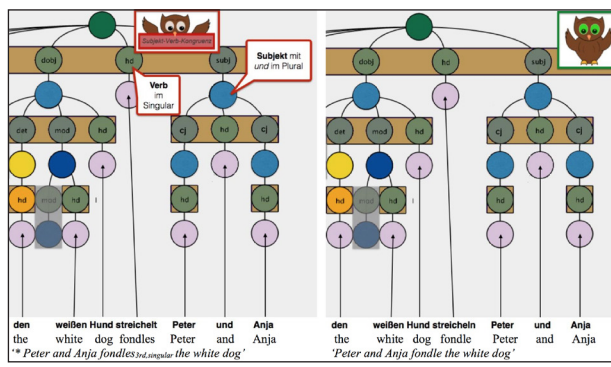
The question we deal with in this paper concerns the format in which to present the feedback. Widely practiced in e-learning systems are animated tutors that interact with the user in a socially engaging manner (for a recent overview, see, e.g. Govindasamy, 2014). Adopting this format, we have chosen an owl – portrayed as intelligent and wise in Aesop's fables – as a character capable to attract and motivate children as well as young adults. The user can select feedback at two levels of detail:

(1) In *verbose mode*, if COMPASS spots an error, the owl shows up at the word or phrase the learner has just attached erroneously, and displays a box with information about the error type (see left panel in Figure 2 for an example of missing subject-verb agreement).

(2) In *concise mode*, the owl has green eyes, sitting in a green box at the upper right corner as long as the user is building grammatically correct structures (right panel in Figure 2). In case of an error the color changes to red.

The user can switch off feedback if desired.

Figure 2. Syntactic structures with feedback. Left image: verbose feedback presented after a subject-verb agreement error. Right image: concise feedback after a correct attachment

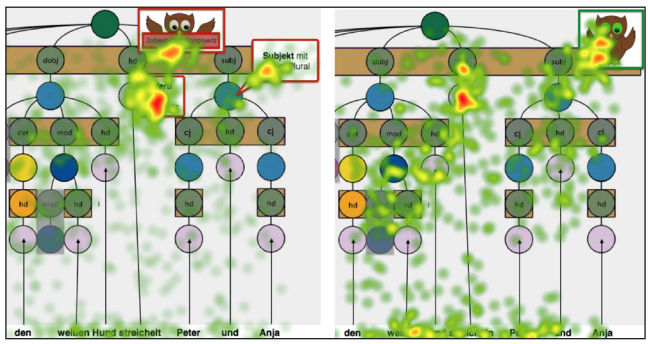


3. A preliminary system evaluation using eye-tracking

In order to find out whether the feedback was sufficiently salient and informative to attract the attention of the users, we performed an eye-tracking experiment with 20 adult native speakers of German (university students) who were instructed to judge the grammatical correctness of displayed sentences. We also asked the participants to express in a questionnaire their subjective impression of the usefulness of the feedback.

The experiment comprised two phases. In Phase I, the participants familiarized themselves with COMPASS by freely composing syntactic trees in drag-and-drop dialogues with the system. Then, in Phase II, they were presented with prefabricated sentences featuring correct or incorrect subject-verb agreement (adapted from an experiment by Bock & Miller, 1991), allowing them to exercise relatively rare and error-prone agreement cases. In addition to these experimental sentences, there were filler sentences that were either grammatically correct or contained other types of errors. In one experimental condition, the participant delivered his/her grammaticality judgment, and the owl indicated the (un)grammaticality of the input sentence in the manner illustrated in Figure 2. All sessions were video-recorded and transcribed.

Figure 3. 'Heat maps' representing fixation durations for the images in Figure 2. Green spots: short fixations; fixation durations increase via yellow to red. Left image: fixation pattern in case of a subject-verb agreement error; right image: fixations in case of positive feedback



As indicated by the 'hot spots' in the heat maps (Figure 3), the participants did pay attention to the feedback agent. The percentage of correct answers increases from 52% in the control condition (no tree, no feedback) to 67% in the experimental

condition with syntactic trees (more precisely, in the experimental subcondition without feedback where the owl did not give away the correct answer). This suggests the feedback was not only perceived but also yielded a learning effect. The questionnaire data showed that several participants (5 out of 20) found the trees baroque and confusing rather than helpful. These participants did not profit from the feedback.

4. Conclusions

The results of our preliminary user evaluation experiment suggest that our way of presenting feedback via an animated pedagogical agent is promising. However, the questionnaire revealed considerable dissatisfaction with the level of detail of the linguistic information provided by COMPASS. The participants' complaints ranged from too many colors to too much linguistic sophistication. We take these results as recommendations to continue with feedback presentation in the form of animated tutors and with syntactic trees, but also to simplify the tree format considerably.

5. Acknowledgements

We would like to thank Kashif Arshad who conducted the experiment. Moreover, we are indebted to Gerard Kempen for valuable comments on earlier versions of the paper.

References

Ballestracci, S. (2005). *Zum DaF-Erwerb ausgewählter grammatischer Strukturen der deutschen Sprache bei italophonen Studierenden der Pisaner Facoltà di Lingue e Letterature Stranieret.* Doctoral dissertation. Università di Pisa, Italy.

Bock, K., & Miller, C. A. (1991). Broken agreement. *Cognitive Psychology, 23*(1), 45-93. https://doi.org/10.1016/0010-0285(91)90003-7

Brants, S., Dipper, S., Eisenberg, P., Hansen-Schirra, S., König, E., Lezius, W., Rohrer, C., Smith, G., & Uszkoreit, H. (2004). TIGER: linguistic interpretation of a German corpus. *Research on Language and Computation, 2*(4)*,* 597-620. https://doi.org/10.1007/s11168-004-7431-3

Diehl, E., Christen, H., Leuenberger, S., Pelvat, I., & Studer, T. (2000). *Grammatikunterricht: Alles für der Katz?* Tübingen: Niemeyer.

Govindasamy, M. K. (2014). Animated pedagogical agents: a review of agent technology software in electronic learning environments. *Journal of Educational Multimedia and Hypermedia, 23*(2), 163-188.

Harbusch, K., Cameran, C.-J., & Härtel, J. (2014). Feedback to grammatical errors in German as second language focused on the learner's personal acquisition level. In S. Jager, L. Bradley, E. J. Meima & S. Thouësny (Eds), *CALL design: principles and practice* (pp. 140-145). Dublin: Research-publishing.net. https://doi.org/10.14705/rpnet.2014.000208

Harbusch, K., Härtel, J., & Cameran, C.-J. (2013). COMPASSIII: teaching L2 grammar graphically on a tablet computer. *Proc. of SLaTE 2013, Grenoble, France*.

Harbusch, K., & Kempen, G. (2011). Automatic online writing support for L2 learners of German through output monitoring by a natural-language paraphrase generator. In M. Levy, F. Blin, C. Bardin Siskin & O. Takeuchi (Eds), *WORLDCALL* (pp. 128-143). New York: Routledge.

Kartchava, E. (2012). *Noticeability of corrective feedback, L2 development and learner beliefs*. PhD Thesis. Université de Montréal, Canada.

Kempen, G., & Harbusch, K. (2002). Performance grammar: a declarative definition. In M. Theune, A. Nijholt & H. Hondorp (Eds), *Computational linguistics in the Netherlands 2001*. Amsterdam: Rodopi. https://doi.org/10.1163/9789004334038_013

Molloy, E. K. & Boud, D. (2014). Feedback models for learning, teaching and performance. In J. M. Spector, M. D. Merrill, J. Elen & M. J. Bishop (Eds), *Handbook of research on educational communications and technology* (pp. 413-424). New York: Springer.

Narciss, S. (2008). Feedback Strategies for Interactive Learning Tasks- In J. M. Spector, M. D. Merrill, J. J. G. van Merriënboer, & M. P. Driscoll (Eds.), *Handbook of Research on Educational Communications and Technology* (pp. 125-144). Mahwah, NJ: LEA

Van der Kleij, F. M. (2013). *Computer-based feedback in formative assessment.* PhD Thesis. University of Twente, Enschede, the Netherlands.

Varnosfadrani, A. D., & Ansari, D. N. (2011). The effectiveness of error correction on the learning of morphological and syntactic features. *World Journal of English Language, 1*(1), 29-40. https://doi.org/10.5430/wjel.v1n1p29

Research-publishing.net

Published by Research-publishing.net, not-for-profit association
Dublin, Ireland; Voillans, France, info@research-publishing.net

CALL communities and culture – short papers from EUROCALL 2016
Edited by Salomi Papadima-Sophocleous, Linda Bradley, and Sylvie Thouësny

Typeset by Research-publishing.net
Cover design by © Easy Conferences, info@easyconferences.eu, www.easyconferences.eu
Cover layout by © Raphaël Savina (raphael@savina.net)
Photo "bridge" on cover by © Andriy Markov/Shutterstock
Photo "frog" on cover by © Fany Savina (fany.savina@gmail.com)
Fonts used are licensed under a SIL Open Font License

ISBN13: 978-1-908416-43-8 (Paperback - Print on demand, black and white)
Print on demand technology is a high-quality, innovative and ecological printing method; with which the book is never 'out of stock' or 'out of print'.

ISBN13: 978-1-908416-44-5 (Ebook, PDF, colour)
ISBN13: 978-1-908416-45-2 (Ebook, EPUB, colour)

Legal deposit, Ireland: The National Library of Ireland, The Library of Trinity College, The Library of the University of Limerick, The Library of Dublin City University, The Library of NUI Cork, The Library of NUI Maynooth, The Library of University College Dublin, The Library of NUI Galway.

Legal deposit, United Kingdom: The British Library.
British Library Cataloguing-in-Publication Data.
A cataloguing record for this book is available from the British Library.

Legal deposit, France: Bibliothèque Nationale de France - Dépôt légal: décembre 2016.